Insights into the degradation capacities of Amycolatopsis tucumanensis DSM 45259 guided by microarray data

Bourguignon, Natalia; Bargiela, Rafael; Rojo, David; Chernikova, Tatyana; Lopez de Rodas, Sara A.; Garcia-Cantalejo, Jesus; Nather, Daniela J.; Golyshin, Peter; Barbas, Coral; Ferrero, Marcelo; Ferrer, Manuel

World Journal of Microbiology and Biotechnology

DOI:
10.1007/s11274-016-2163-8

Published: 01/12/2016

Peer reviewed version

Journal: *World Journal Microbiology and Biotechnology*

Insights into the degradation capacities of *Amycolatopsis tucumanensis* DSM 45259 guided by microarray data

Natalia Bourguignon[1,#], Rafael Bargiela[2,#], David Rojo[3#], Tatyana N. Chernikova[4], Sara A López de Rodas[5], Jesús García-Cantalejo[5], Daniela J Näther[6], Peter N Golyshin[4], Coral Barbas[3], Marcela Ferrero[1,§], and Manuel Ferrer[2,§*]

[1] Planta Piloto de Procesos Industriales Microbiológicos (PROIMI-CONICET), Tucumán, Argentina.

[2] Consejo Superior de Investigaciones Científicas (CSIC), Institute of Catalysis, Madrid, Spain.

[3] Centro de Metabolómica y Bioanálisis (CEMBIO), Facultad de Farmacia, Universidad CEU San Pablo, Campus Monteprincipe, Boadilla del Monte, Madrid, Spain.

[4] School of Biological Sciences, Bangor University, Gwynedd LL57 2UW, UK.

[5] Unidad de Genómica-Campus Moncloa, C.A.I. Genómica y Proteómica, Facultad CC. Biológicas, Universidad Complutense de Madrid, Madrid, Spain.

[6] Institute for Microbiology, Goethe University, Biocentre, Frankfurt, Germany.

***Correspondence**

Manuel Ferrer

mferrer@icp.csic.es

[#]These authors contributed equally to this work

[§]These authors contributed equally to this work

Running title: Catabolome of *A. tucumanensis*

Abstract The analysis of catabolic capacities of microorganisms is currently often achieved by cultivation approaches and by the analysis of genomic or metagenomic datasets. Recently, a microarray system designed from curated key aromatic catabolic gene families and key alkane degradation genes was designed. The collection of genes in the microarray can be exploited to indicate whether a given microbe or microbial community is likely to be functionally connected with certain degradative phenotypes, without previous knowledge of genome data. Herein, this microarray was applied to capture new insights into the catabolic capacities of copper-resistant actinomycete *Amycolatopsis tucumanensis* DSM 45259. The array data support the presumptive ability of the DSM 45259 strain to utilize single alkanes (*n*-decane and *n*-tetradecane) and aromatics such as benzoate, phthalate and phenol as sole carbon sources, which was experimentally validated by cultivation and mass spectrometry. Interestingly, while in strain DSM 45259 *alk*B gene encoding an alkane hydroxylase is most likely highly similar to that found in other actinomycetes, the genes encoding benzoate 1,2-dioxygenase, phthalate 4,5-dioxygenase and phenol hydroxylase were homologous to proteobacterial genes. This suggests that strain DSM 45259 contains catabolic genes distantly related to those found in other actinomycetes. Together, this study not only provided new insight into the catabolic abilities of strain DSM 45259, but also suggests that this strain contains genes uncommon within actinomycetes.

Keywords Alkanes · *Amycolatopsis tucumanensis* · aromatics · catabolome · degradation · microarray

Introduction

Imagine the microbial communities responding to pollutants-intake and how variable this can be (Hazen et al. 2010; Kostka et al. 2011; Beazley et al. 2012; Guazzaroni et al. 2013; Gutierrez et al. 2013). The most obvious reaction will be their ability to react by degrading such pollutants to intermediates feeding the central metabolism (Liang et al. 2011; Lu et al. 2012; Mason et al. 2012; Kimes et al. 2013; Mason et al. 2014). There is great interest in identifying next-generation information that allows predicting the diversity of pollutants that each community and the microorganisms conforming it, can degrade and the catabolic genes implicated (Pérez-Pantoja et al. 2008; Pérez-Pantoja et al. 2012; Guazzaroni et al. 2013; Bargiela et al. 2015a). The analysis of catabolic capacities of microbial communities or single cultures begins by assessing gene contents, which are currently often achieved using genomic or metagenomic data (Guazzaroni et al. 2013; Bargiela et al. 2015a,b), followed by the analysis of the annotated genome or metagenome and a reference catabolic database as input information (Pérez-Pantoja et al. 2008; Pérez-Pantoja et al. 2012; Guazzaroni et al. 2013; Bargiela et al. 2015a,b). Further, catabolic network can be built, using as an input potential protein-coding gene sequences obtained by direct sequencing of DNA material and the web-based AromaDeg resource (Duarte et al. 2014; Bargiela et al. 2015a,b).

In case genome information is lacking, the identification of catabolic capacities required extensive experimental efforts, i.e. by producing microcosms in which the ability to degrade pollutants is investigated using labeled or not labeled compounds (Watanabe and Hamamura 2003; Pandey et al. 2008). With the aim of easing this process, a novel internally calibrated functional gene microarray system (the so-called catabolome array) was recently developed (Vilchez-Vargas et al. 2013). It contains optimally designed probes covering key aromatic catabolic gene families and key

alkane degradation genes. This enables identifying molecular functions of identified genes in light of catabolic pathways by using DNA material, without the need of genome sequencing. The microarray contains 3605 probes (50 mer) representing catabolic gene subfamilies encoding key activities in hydrocarbon degradation pathways, that included Rieske non-haem iron ring hydroxylating (di)oxygenases (RHDO), extradiol dioxygenases of the vicinal chelate superfamily (EXDOI), intradiol dioxygenases (INDO), soluble di-iron aromatic ring hydroxylating monooxygenases, ferredoxins of multicomponent aromatic degradation enzymes (FERRE), muconate cycloisomerases (MCIS), maleylacetate reductases (MACR), alkane hydroxylases of the integral membrane-bound monooxygenases (ALKB), cytochrome P450, CYP153 alkane hydroxylases (CYP153), benzoyl coenzyme A reductases (BCOAR), and benzylsuccinate synthases. Supplementary Table 1 provides information regarding accession numbers for sequences and taxonomic origin of catabolic genes associated to each of the probes. Briefly, most probes (circa 84%) derived from genomes from cultivable bacteria of at least 182 different species, distributed among 70 genera that included *Gordonia*, *Nocardioides*, *Rhodococcus*, *Prauserella*, *Mycobacterium*, *Nocardia*, *Dietzia*, *Corynebacterium*, *Frankia* and *Janibacter* (Actinobacteria), *Flavobacteria*, *Dokdonia*, *Polaribacter* and *Maribacter* (Bacteroidetes), *Geobacillus* and *Desulfitobacterium* (Proteobacteria), *Acinetobacter*, *Sphingomonas*, *Alcanivorax*, *Cycloclasticus*, *Pseudomonas*, *Legionella*, *Xanthomonas*, *Burkholderia*, *Oleiphilus*, *Xanthobacter*, *Thalassolituus*, *Acidisphaera*, *Photorhabdus*, *Bdellovibrio*, *Ruegeria*, *Rhodobacter*, *Ralstonia*, *Methylococcus*, *Bradyrhizobium*, *Hahella*, *Jannaschia*, *Polaromonas*, *Paraburkholderia*, *Paracoccus*, *Marinobacter*, *Sulfitobacter*, *Roseovarius*, *Oceanicola*, *Pseudooceanicola*, *Oceanicaulis*, *Loktanella*, *Maritimibacter*, *Parvularcula*, *Roseobacter*, *Acidiphilium*, *Psychrobacter*, *Bermanella*,

Stenotrophomonas, *Blastochloris*, *Azoarcus*, *Magnetospirillum*, *Geobacter*, *Thauera*, *Ensifer*, *Aromatoleum*, *Rhodopseudomonas*, *Syntrophobacter*, *Alkalilimnicola*, *Desulfobacula*, *Parvibaculum*, *Sphingopyxis*, *Caulobacter*, *Erythrobacter* and *Novosphingobium* (Proteobacteria). Probes from species of *Tetrahymena* (Eukaryotia), *Neurospora* (Ascomycota) and *Methanopyrus* (Euryarchaeota) are also included. Note that within bacterial species whose probes are included in the microarray, 30 belong to 10 genera (*Gordonia*, *Nocardioides*, *Rhodococcus*, *Prauserella*, *Mycobacterium*, *Nocardia*, *Dietzia*, *Corynebacterium*, *Frankia* and *Janibacter*) of the order Actinomycetales. Detailed information on all the probes on the array and the evolutionary relationships are reported elsewhere (Vilchez-Vargas et al. 2013).

In this work, we exploit the catabolome array (Vilchez-Vargas et al. 2013) to get new insights into the degrading capacities of *A. tucumanensis* strain DSM 45259, a copper-resistant actinobacterium isolated from polluted sediments (Albarracín et al. 2010a). Note that the threshold for considering a signal as a true positive in the array was set when hybridization occurred with a probe exhibiting > 80% sequence identity, where it can be assumed that the target DNA is derived from a gene encoding a member of the same subfamily as that for which the probe was designed (Vilchez-Vargas et al. 2013). This, together with the fact that the majority of the probes belong to bacteria, including actinomycetes (see above), suggest that catabolic genes of the actinomycete *Amycolatopsis tucumanensis* DSM 45259 will be detectable. Having said that, *A. tucumanensis* DSM 45259 was widely studied for its remarkable copper-resistance (Dávila Costa et al. 2011a,b; Dávila Costa et al. 2012). More recently, degradation of naphthalene and phenanthrene was found to occur in minimal medium when growing on glucose as co-substrate (Bourguignon et al. 2014). In the present study we provide evidences that *A. tucumanensis* DSM 45259 has also the capacity to use aliphatic and

aromatic hydrocarbons such as *n*-decane, *n*-tetradecane, phthalate, benzoate and phenol as sole carbon sources. These abilities, predicted by the microarray, were further confirmed by cultivation tests and target mass spectrometry analysis. Although, such degradation capacities are common within other actinomycetes, the results suggest that strain DSM 45259 carries some catabolic genes distantly related to previous catabolic genes of other actinomycetes. In addition to that, because good agreement with the array-based predictions was observed after experimental validations, we suggest that the strategy herein described represents a promising strategy for disentangling contexts-specific catabolic phenotypes in any organism or microbial community, without the need of genome or metagenome sequencing.

Materials and Methods

Chemicals and basic culture conditions

All chemicals used were of the purest grade available and were purchased from Fluka-Aldrich-Sigma Chemical Co. (St Louis, MO, USA). *A. tucumanensis* strain DSM 45259, a copper resistant strain, was used in this study (Albarracín et al. 2010a; Dávila Costa et al. 2012). Strain DSM 45259 was cultivated in Tryptic Soy Broth (TSB) medium (tryptein: 17 g L^{-1}; soy peptone: 3 g L^{-1}; NaCl: 5 g L^{-1}; K_2HPO_4: 2.5 g L^{-1}; glucose: 2.5 g L^{-1}; pH 7.3 ± 0.2) at 30 °C until late exponential growth phase. This culture was used as pre-inoculum for a 30 ml Minimal Media (MM) broth ($(NH_4)_2SO_4$: 2 g L^{-1}; K_2HPO_4: 0.5 g L^{-1}; $MgSO_4.7H_2O$: 0.2 g L^{-1}; $FeSO_4$: 0.01 g L^{-1}; glucose: 1.25 g L^{-1}; pH 7.0 ± 0.2) containing 0.2 mM naphthalene (from a 25 mM stock solution in acetone) and glucose 1.25 g L^{-1}. Control culture without the addition of hydrocarbon was performed. Cultures were incubated at 30 °C and 180 rpm for 96 h, after which

147 cultures were centrifuged (8000 rpm; 10 min; 4 °C) and cell pellets used for DNA
148 extraction.

149 **Catabolome microarray analysis**

150 The total DNA extraction was done by using cells harvested during late exponential
151 growth phase in cultures containing glucose and naphtalene as carbon sources, and the
152 hexadecyltrimethylammonium bromide (CTAB) method with some modifications
153 (Bailey et al. 1995). Briefly, harvested cells were re-suspended in 750 µL lysozyme-
154 CTAB extraction solution (8 mg mL^{-1} lysozyme, 2 % CTAB, 1.4 M NaCl, 20 mM
155 ethylenediaminetetraacetic acid (EDTA), 100 mM Tris-HCl, pH 8, 50 mg L^{-1} ARNase,
156 0.3 M sucrose). After incubation during 2 h at 37 ºC to improve cell lysis, 250 µL
157 sodium dodecyl sulfate (SDS) 2% (w/v) were added, the solution was vortexed for 1
158 min, and then 2 µL β-mercaptoethanol added and incubated 30 min at 60 °C. To purify
159 DNA, 1 volume of chloroform:isoamyl alcohol (24:1) was added. The solution was
160 mixed and centrifuged (12000 rpm, 15 min). After separation of the aqueous phase, 1
161 volume of 2-propanol was added and the solution incubated at -20 °C during 1 h to
162 facilitate DNA precipitation. The precipitated DNA was washed with 1 volume of 70 %
163 (v/v) ethanol and dried. Finally, the DNA was re-suspended in 50 µL sterile distilled
164 water. Purity of extracted DNA was assessed by measuring the 260/280 and 260/230
165 ratios using a spectrophotometer. DNA concentration was measured using Quanti-iT
166 dsDNA Assay kit (Invitrogen, Paisley, UK).

167 For DNA fragmentation, the resulting genomic DNA was heat-fragmented at 95 °C
168 for up to 1.5 h. The aliquot of each digestion reactions was analyzed on 2% (w/v)
169 agarose gel and completed if the majority of DNA fragments had a size range of 200-
170 1000 base pairs (bp). This resulted in the production of fragments 200 to 1000 bases in

length. The resulted DNA was precipitated with isopropanol, suspended in 45 µL of MilliQ water and used for labelling. We labelled total DNA by direct incorporation of Cy5-conjugated dUTP (GE Healthcare) using terminal deoxynucleotidyl transferase (Thermo Scientific, Paisley, UK). Following a 4 h incubation at 37 ºC, the reaction terminated by addition of 0.5 M EDTA, pH 8.0. The labelled target was purified from unincorporated dye molecules by adding 200 µL of TE buffer and spinning through a Microcone filter (Millipore, Hertfordshire, UK) for 15 minutes at 11000 rpm. The purified, labelled target was precipitated with isopropanol, and resuspended to a final volume of 20 µL with MilliQ water. The dye incorporation was measured with a NanoDrop spectrophotometer. Labelled DNA samples were vacuum-dried and stored at -20º C until hybridization.

We used a chip previously designed and calibrated by Vilchez-Vargas et al. (2013). For microarray hybridization, probes were printed on CodeLink Activated slides (SurModics, Eden Prairie, USA) using MicroGrid TAS II spotter (BioRobotics, Germany) at the University of Frankfurt (Frankfurt, Germany). Coupling of DNA probes was performed by overnight incubation of slides in saturated NaCl chamber. Post-coupling processing included the blocking of residual reactive groups and was done as follows: slides were washed with 4x SSC (190 mM sodium chloride plus 20 mM tri-sodium citrate equivalent to sodium concentration of 250 mM), 0.1% (w/v) SDS, for 30 min, then rinsed briefly with deionized water and dried by centrifugation for 3 min at low-speed centrifuge. Prior to hybridization, labelled DNA was incubated with herring sperm DNA (Invitrogen, Paisley, UK) for 5 min at 95 °C and then 80 µL of hybridization buffer was added. For hybridization, slides were inserted into hybridization chamber and after that were covered by coverslips. The solution of Cy5 –

dUTP labelled DNA in hybridization buffer (100 µL total volume) was carefully infused through narrow gaps between slides and covers.

The hybridization was performed at 55 °C for 18 h using hybridization buffer consisting of 15% (v/v) dimethylsulfoxide, 25% (v/v) formamide, 1.25 x SSC, 0.15% (w/v) SDS, 0.15% (w/v) Tween 20, 880 mM betaine, 5x TE buffer (50 mM Tris-HCl, 5 mM EDTA) and 0.1 mg L^{-1} bovine serum albumin (BSA) in aqueous solution. Following hybridization, slides were washed 5 min at 42 ºC in 1x SSC containing 0.3% (w/v) SDS, twice in 1x SSC (1 min, 42 ºC), in 0.5x SSC (1 min, 20 °C), in 0.1x SSC containing 0.3% (w/v) SDS (1 min, 42° C) and finally twice in 0.1x SSC (1 min, 20 °C). Slides were dried at low speed in centrifuge for 30 seconds. Slides were scanned in a GenePixR 4000B microarray scanner (Molecular Devices, Berkshire, UK) and images analyzed by using the software of image analysis GenePixRPro 6.0 from Axon Instruments / Molecular Devices Corp (Molecular Devices, Berkshire, UK).

Each query sequence from probes targeting catabolic genes (see accession numbers in Supplementary Table S1) for which a positive signal was obtained in the microarray was submitted to web-based AromaDeg resource (Duarte et al. 2014). Each sequence was then associated with a catabolic enzyme performing an aromatic compound degradation reaction.

Quantification of the biodegradation of hydrocarbons by cultivation and liquid chromatography-mass spectrometry

Activation of the DSM 45259 strain was firstly done by transfers in minimal-mineral medium with a low concentration of hydrocarbon (20 mg L^{-1}) for the adaptation of the microorganism. The medium consisted of 2.6 g Na_2HPO_4, 1.33 g KH_2PO_4, 1 g $(NH_4)_2SO_4$ and 0.20 g $MgSO_4 \cdot 7\ H_2O$ dissolved in 1000 mL of demineralized water. The

medium was adjusted to pH 7.2 ± 0.3. After sterilization, 5 mL of trace element solution and 1 mL of vitamin solution were added. Both solutions were prepared as described in DSMZ methanogenium medium 141 and autoclaved or sterile filtered separately ([DSMZ 2012]). Substrates were sterilized separately and added aseptically at an amount of 20 mg L^{-1} each. Cultivation was done at 30 °C with a 180 rpm constant agitation during 72 h. The cell biomass was washed twice with 20 mM sodium phosphate buffer pH 7.0 and used to produce cultures. Briefly, cell pellets were grown in 30 mL of the same medium (0.4 g wet cell pellet L^{-1}) with various aliphatic and aromatic hydrocarbons such as *n*-decane, *n*-tetradecane, phenol, benzoate and phthalate to serve as the sole source of carbon and energy. Substrates were sterilized separately and added aseptically at an amount of 500 mg L^{-1} each. Cultivation was done at 30 °C with a 180 rpm constant agitation during 72 h. Two controls were done: a control test without the addition of the cells (abiotic test) and a control test without the addition of the hydrocarbon (biotic test).

The extraction of the hydrocarbons and their degradation intermediates was performed by adding 1 volume of acetone to the cultures. After homogenization, flasks were stand for 30 min, and then centrifuged at 13000 rpm during 10 min. The supernatants were analyzed by target analysis by Liquid Chromatography (LC)-Mass Spectrometry (MS) to confirm the degradation of the initial substrates as well as the existence of degradation intermediates in test and control cultures. For that, the following reagents and standards have been used: acetonitrile (LC-MS grade, Sigma-Aldrich, Steinheim, Germany), formic acid (FA) (MS grade, Sigma-Aldrich, Steinheim, Germany) and MilliQ® water (Millipore, Billerica, MA, USA). For reference masses purine, hexakis(1H,1H,3H-tetrafluoropropoxy)phosphazine (HP) and ammonium trifluoroacetate ($TFA(NH_4)$) from Agilent (API-TOF reference mass solution kit) were

used. The metabolic profile was achieved using a liquid chromatography system consisting of a degasser, a binary pump, and an auto-sampler (1290 infinity II, Agilent). Samples (0.5 µL) were applied to a reversed-phase column (Zorbax Extend C18 50 x 2.1 mm, 1.8 µm; Agilent), which was maintained at 60 °C during the analysis. The system was operated at a flow rate of 0.6 mL min^{-1} with solvent A (water containing 0.1% formic acid) and solvent B (acetonitrile containing 0.1% formic acid). The gradient was 5% B (0–1 min), 5 to 80% B (1–7 min), 80 to 100% B (7–11.5 min), and 100 to 5% B (11.5–12 min). The system was finally held at 5% B for 3 min to re-equilibrate the system (15 min of total analysis time). Data were collected in positive and negative ESI modes in separate runs using QTOF (Agilent 6550 iFunnel). The analyses were performed in both positive and negative ion modes in full-scan from *m*/*z* 50 to 1000. The capillary voltage was 3000 V and the nozzle voltage was 1000 V with a scan rate of 1.0 spectrum per second. The gas temperature was 250 °C, the drying gas flow was 12 L min^{-1}, the nebulizer was 52 psi, the sheath gas temperature 370 °C and the sheath gas flow 11 L min^{-1}. For positive mode, the MS-TOF parameters were as follows: fragmentor 175 V and octopole radio frequency voltage 750 V. For negative mode, the MS-TOF parameters included the following: fragmentor 250 V and octopole radio frequency voltage 750 V. During the analyses, two reference masses were used: 121.0509 (purine, detected *m*/*z* $[C_5H_4N_4+H]^+$) and 922.0098 (HP, detected *m*/*z* $[C_{18}H_{18}O_6N3P_3F_{24}+H]^+$) in positive mode and 112.9856 (TFA(NH_4), detected *m*/*z* $[C_2O_2F_3(NH_4)-H]^-$) and 966.0007 (HP+FA, detected *m*/*z* $[C_{18}H_{18}O_6N_3P_3F_{24}+FA-H]^-$) in negative mode. The references were continuously infused into the system, enabling constant mass correction. Samples were analyzed in randomized runs, during which they were incubated in an auto-sampler at 4 °C. The analytical runs for both polarities

were set up starting with the analysis of ten equilibrium injections followed by the samples. A single injection per sample was done.

Based on a list of candidates, their accurate monoisotopic masses were searched for in the MS chromatograms (± 10 ppm) using MassHunter Quantitative Analysis (B.06.00, Agilent) and their identification confirmed by the analysis of the commercial standards. Then, the corresponding peak areas were integrated using the same software.

Determination of catechol dioxygenase activity

To investigate the catechol 2,3-dioxygenase activity, a spectrophotometric method was used, in which the formation of oxidation products is followed. Briefly, the strain was pre-cultivated in minimal-mineral medium as described before at 30 ºC until exponential phase using benzoate and phenol as sole carbon sources (20 mg L^{-1}). This culture was used to inoculate 30 mL of minimal-mineral medium containing 500 mg L^{-1} of benzoate and phenol, respectively (for details see above). Control cultures without the addition of hydrocarbons were done. Cultures were incubated at 30 ºC at 180 rpm during 72 h, after which cells were separated by centrifugation (8000 rpm; 10 min; 4 ºC). The pellet was washed twice with 50 mM K/Na-phosphate (pH 7.5) buffer, and then re-suspended in 5 mL of this buffer. For the preparation of protein cell extracts, the cells were broken by three passages in a French Press® at 20000 psi, after which the sample was centrifuged (10000 rpm; 10 min; 4 °C) to eliminate cell debris. Supernatant was carefully aspirated and immediately used for activity assay. The assay was performed in 96-well plates and 200 µL of total volume, as described elsewhere (Alcaide et al. 2013). Briefly, the catechol 2,3-dioxygenase activity was measured (in triplicates) in a microplate reader (Synergy HT Multi-Mode Microplate Reader - BioTek) by evaluating the increase of absorbance at 388 nm due to the formation of the

reaction product 2-hydroxy-6-oxohepta-2,4-dienoate (HOHD), in a reaction mixture that contains 10 μL of protein extract containing catechol 2,3-dioxygenase to a substrate in the presence of the following solution: 87 μL of K/Na-phosphate (pH 7.5) and 3 μL of catechol solution in H_2O (10 mM) to achieve a final substrate concentration of 0.15 mM. Reactions were followed at 30 ºC for 20 min (ε_{HOHD} at 388 nm = 13,800 M^{-1} cm^{-1}). One unit (U) of enzyme activity was defined as the amount of enzyme required to transform 1 μmol of substrate in 1 min under the assay conditions.

Results

Degradation capacities of DSM45259 guided by microarray data

Catabolome microarray data from *A. tucumanensis* DSM 45259, grown in napthalene and glucose as co-substrate constituted the input information in our study. The complete information about the microarray tests is described in the Materials and Methods section. Following the restrictive criterion of fold-change above 6-fold higher than background signal in the internally calibrated microarray system, we detect a total of 5 out of 3605 genes encoding proteins with proved catabolic functions (Table 1). As mentioned, to detect signals with a high precision, only signals > 6 normalized intensity (NI) were considered, as the use of internal positive controls for setting the correct threshold according to the desired precision of the experiment revealed that, under conditions described in Materials and Methods, any signal > 6-8NI is highly unlikely to be false positive (Vilchez-Vargas et al. 2013). Raw fluorescence signals for probes targeting the 3605 genes are detailed in Supplementary Table 1. Within the 11 catabolic genes families targeted by the microarray (Vilchez-Vargas et al. 2013), we detected the presence of 4 covered by the 5 positive probes, which are summarized below.

315 Genes implicated in alkane degradation were found. Particularly, the probes
316 targeting *alk*B genes AJ833983 and AJ833926 for AlkB alkane hydroxylases (the so-
317 called ALKB catabolic gene family by Vilchez-Vargas et al. 2013) were strongly
318 detected (Table 1). AlkB participates in the initial attack of n-alkenes in the *n*-alkane
319 oxidation pathway (Fig. 1). Within ring hydroxylating dioxygenases (RHDO catabolic
320 gene family) we found AAD17377 by a high intensity of hybridization, followed by
321 AAD03558 (Table 1). According to AromaDeg (Duarte et al. 2014), AAD17377 gene
322 encodes a benzoate 1,2-dioxygenase (Bzt) that convert benzoate into *cis*-1,6-dihydroxy-
323 2,4-cyclohexadiene-1-carboxylic acid within the benzoate to catechol degradation
324 pathway, and AAD03558 a phthalate 4,5- dioxygenase (Pht) that converts phthalate into
325 protocatechuate. Finally, within ring hydroxylating monooxygenases (RHMO catabolic
326 gene family) we found Z36909, which was the probe with the highest level of
327 hybridization intensity (Table 1), and that encodes a phenol hydroxylase, an enzyme
328 that catalyzes the first step in the degradation of phenol into catechol.

329 Taken together, the microarray data support the ability of the DSM 45259 strain to
330 utilize single alkanes and aromatics such as benzoate and phenol (through conversion to
331 catechol) and phthalate (through conversion to protocatechuate) as carbon sources (Fig.
332 1). Interestingly *alk*B gene in the strain DSM 45259 matches with two probes
333 (AJ833983 and AJ833926) encoding the same protein, namely, an alkane hydroxylase
334 (AlkB) from actinomycete *Rhodococcus* species; this matches with the taxonomy of
335 strain DSM 45259. Indeed, several actinomycetes able to degrade C_5–C_{10} alkanes
336 contain alkane hydroxylases as, for example, representatives from mycobacteria and
337 rhodococci (van Beilen et al. 2005; Sekine et al. 2006; Lincoln et al. 2015). By contrast,
338 genes encoding Bzt (AAD17377) and Pht (AAD03558) match with probes from
339 Proteobacteria (*Sphingobium* and *Burkholderia* spp.), and that of the phenol

hydroxylase (Z36909) to a probe from *Acinetobacter* sp. This suggests that *alk*B gene in DSM 45259 strain is highly similar to that found in other actinomycetes, while the other 3 genes are quite divergent to those from actinomycetes.

Note that no any other gene implicated in the later stages of the degradation of alkanes, apart from *alk*B, was detected in the microarray because it does not contains such genes (Vilchez-Vargas et al. 2013). In the case of genes implicated in the later stages of the degradation of protocatechuate, phenol and catechol, the microarray contains probes encoding catechol-2,3-dioxygenases (Cat) within the so-called extradiol dioxygenases (EXDO) catabolic gene family, and catechol 1,2-dioxygenases and protocatechuate 3,4-dioxygenases (3,4-PCD) within the so-called intradiol dioxygenase (INDO) catabolic gene family. None of those genes were detected in the microarray according to 6 fold-change criterion, suggesting the absence or low expression level of those genes in strain DSM 45259 under our assay conditions. Indeed, the DNA material used for microarray hybridization was obtained from cells harvested during late exponential growth phase in cultures containing glucose and naphtalene as carbon sources, where those genes may be expressed at low level. This is in agreement with cultivation, activity tests and mass spectrometry experiments (see below) that confirmed that strain DSM 45259 contains 3,4-PCD and Cat activities when grown on phthalate and benzoate (see below), and that in the absence of these substrates expression level of those genes may be most likely low. This is not the case of the genes encoding catechol 1,2-dioxygenase whose presence in the genome of strain DSM 45259 could not be confirmed both by array and cultivation tests (see below).

Experimental validation by cultivation and mass spectrometry

To prove the correctness of the predictions and to discard that the predictions are an artifact derived from an inaccurate hybridization, experimental validation assays were conducted. For that, cultures were set up with C_{10} and C_{14} alkanes (*n*-decane and *n*-tetradecane), and the aromatics phthalate, phenol and benzoate as the only carbon source, and after 0, 24, 48 and 72 h cultivation we examined the efficiency of strain DSM 45259 to degrade them. A concentration of 500 mg L^{-1} of each compound was used. Target analysis by Liquid Chromatography-Mass Spectrometry (LC-MS) was further used to confirm the consumption of the initial substrates and the formation of key degradation intermediates in test cultures as compared to the abiotic (culture media containing aromatics but no cells) and biotic (culture without the aromatics) control cultures.

The level of degradation of *n*-decane and *n*-tetradecane could not be obtained as both chemicals could not be detected under our analytical platform. However, *n*-decanoic (in *n*-decane microcosm) and *n*-tetradecanoic acid (in the *n*-tetradecane microcosm) were detected at high level (Table 2), demonstrating that the degradation of both alkanes by strain DSM 45259 occurred. Degradation of phthalate, phenol and benzoate was achieved at 49.7, 89.1 and 57.6%, respectively, at the end of the 72 h assay. This was shown by measuring the remaining amount of these 3 compounds (Table 2). Degradation was further demonstrated by identifying the increasing abundance level of the phthalic-degradation products protocatechuic acid and 3-oxoadipic acid (in phthalate microcosm) and the phenol- and benzoate-degradation product catechol (in phenol and benzoate microcosms) during the follow-up assay (Table 2). The identification of 3-oxoadipic acid in the cultures grown with phthalate in combination to the identification of a gene encoding a phthalate 4,5-dioxygenase in the microarray demonstrate that catabolism of phthalate proceeds via the proto-catechuic

388 *ortho* cleavage pathway in which a protocatechuate 3,4-dioxygenase may be implicated
389 (Fig. 1). In case of catechol degradation, no intermediates were detected above the
390 detection limit by LC-MS, possibly because they are rapidly converted and thus
391 accumulated at low level under cultivation conditions. However, the demonstration of
392 catechol 2,3-dioxygenase activity in protein extracts from cells obtained in cultures
393 grown on benzoate and phenol in MM broth (see Materials and Methods section)
394 revealed that the catechol *meta*-ring cleavage branch is fully operative in DSM 45259
395 (Fig. 1). Indeed, activity values of 0.86 ± 0.07 and 1.18 ± 0.05 unit mg^{-1} protein were
396 obtained under our experimental assay conditions.

397 Taken together, as shown in Table 2, signatures for the degradation of the 5
398 chemicals predicted as being used as carbon sources (Fig. 1) were experimentally found
399 (Table 2), thus confirming a total agreement with our predictions.

400 **Discussion**

401 In this report, we described new insights into the degradation capacities of the copper-
402 resistant actinomycete *A. tucumanensis* DSM 45259 using microarray data. Our
403 approach was based on the utilization of the catabolome microarray presented by
404 Vilchez-Vargas et al. (2013). However, we adapt the output of the microarray data to
405 incorporate a prediction tool based on the utilization of the web-based AromaDeg
406 resource (Duarte et al. 2014), and identify unambiguously genes encoding catabolic
407 proteins of this microorganism. Further, with cultivation and metabolomics approaches
408 being developed, we provided experimental validation. Based on the microarray and
409 experimental data presented we unambiguously identified that *A. tucumanensis* DSM
410 45259 has the ability to use alkanes (i.e., *n*-decane and *n*-tetradecane), phthalate, phenol
411 and benzoate as sole carbon sources. Degradation occurred in the absence of glucose as

co-substrate that was previously reported to be required for the degradation of naphthalene and phenanthrene (Bourguignon et al. 2014).

Actinomycetes possess potent capacities to metabolize aliphatic and aromatic toxic hydrocarbons. Species of the genera *Mycobacterium*, *Streptomyces* and *Nocardia* (that contain genera such as *Gordonia* and *Rhodococcus*), commonly found in contaminated soils, are the best characterized members. Thus, by using cultivation, phylogenetic, phenotypic, and/or genomic information approaches, members of these genera have been shown to use as sole carbon and energy sources, to different extend, a wide range of compounds. They include, crude oil, diesel oil, rapeseed oil, linear and branched medium-to-long chain alkanes (up to C_{36}), alkenes, haloalkanes, monocyclic aromatic compounds (benzoate, catechol, gentisate, salicylate, phenol, phenylethanol, thymol, alkylbenzenes, xylene, toluene, phthalate) and poly-aromatic compounds (biphenyl, naphthalene, anthracene, fluoranthrene, phenanthrene, coronene, pyrene, chrysene, naphthacene, acenapthene, benzo[a]pyrene, 7,12-dimethylbenz[a]anthracene), as well as organic sulfur compounds (i.e. benzothiophene and dibenzothiophene) and nitro-aromatics (Kelley and Cernigilia, 1995; Lloyd-Jones and Hunter, 1997; Bastiaens et al. 2000; Monticello, 2000; Stingley et al. 2004; Kumar et al. 2006; Larkin et al. 2006; Zeinali et al. 2007; Yang et al. 2011; Balachandran et al. 2012; Fathepure, 2014; Luo et al. 2014; Sudhir et al. 2014). In some case biodegradation was only observed when growing on glucose as co-substrate (Pizzul et al. 2006). Members of the genus *Arthrobacter* and *Streptomyces* can also degrade halogenated pesticides (Bourguignon et al. 2014). *Terrabacter* isolates have been also shown to degrade dibenzofuran and the heterocyclic nitrogen compound carbazole (Iida et al. 2002). Therefore, we can conclude that the degrading capacities herein reported for *A. tucumanensis* DSM 45259 are within common abilities for other actinomycetes.

A solid basis of genomic understanding on degradation capacities of actinomycetes has been mainly established, particularly for *Mycobacterium* and *Nocardia* isolates. Thus, genome analysis of *Mycobacteria* isolates has contributed to the characterization of key enzymes such as the initial ring-hydroxylating dioxygenases participating in the degradation of substrates such as biphenyl, naphthalene, anthracene, fluoranthene, pyrene, phenanthrene, phthalate and benzoate (Brezna et al. 2003; Stingley et al. 2004; Kim et al. 2006; Kim et al. 2007; Kim et al. 2008; Kallimanis et al. 2011; Zhang et al. 2012; Kwak et al. 2014). Genome sequence of *Rhodococcus* strains revealed they contain not only multiple alkane hydroxylase genes (*alk*B) and from 27 to 73 cytochrome P450 monooxygenases and other catabolic genes predicted to be involved in the metabolism of alkanes and nitroalkanes, as well as an array of cyclic ketones, halogenated aromatics and aromatic hydrocarbons (e.g., benzoate, catechol, gentisate, salicylate, homogentisate, naphthalene, phenanthrene, anthracene, and benzo[a]pyrene) (McLeod et al. 2006; Chen et al. 2013; Pathak et al. 2013; Zhang et al. 2014; Lincoln et al. 2015; Qu et al. 2015). In addition to that, phylogenetic, phenotypic, and genomic information for 27 completely genome-sequenced mycobacteria revealed a total of 9532 genes conforming the so-called "PAH-degrading" node, of which 3533 genes belong to the core-genome that is present in each strain and 5999 genes belong to the dispensable genome that is absent in one or more strains (Kweon et al., 2015). Among the 3533 core genes, only 136 common genes were tentatively identified to be involved in the degradation of aromatic hydrocarbons, which indicate the high variability of gene sequences and degradation abilities within isolates. Some of these common genes, such as the ones needed for pyrene degradation, have been demonstrated to be acquired by horizontal transfer (DeBruyn et al. 2012). This high genomic variability was also supported by the present study which suggests that at least 3 genes from strain DSM

45259 (those encoding Bzt, Pht and phenol hydroxylase) are quite divergent (<80% sequence identity) to those of other actinomycetes. This was suggested as no hybridization signals of such genes were found with any actinomycete-probes of the same subfamily present in the microarray, while hybridizing with those from Proteobacteria.

In conclusion, we report here new insights into the catabolic abilities of the first member of the *Amycolatopsis* genus and identify a variety of genomic signatures which seems to be uncommon within actinomycetes. This work also contributed to deepening into the degradation capacities of actinomycetes, whose knowledge is mostly limited for *Mycobacterium* and *Nocardia* isolates. Note that the genus *Amycolatopsis* has been classified in the family Pseudonocardiaceae and it currently contains 39 species with validly published names (http://www.bacterio.cict.fr/a/amycolatopsis.html). Recent studies indicate that the chemotaxonomic characteristics of this genus, which relate to, but differentiate from *Streptomyces* and *Nocardia*, are intrinsically determined by the molecular phylogeny of their encoding genes (Xu et al. 2014). Its closest relatives are *Amycolatopsis* sp. ATCC 39116, *A. methanolica* 239 and *A. thermoflava* N1165. *Amycolatopsis* sp. ATCC 39116 (previously known as *S. setonii*) harbors genes encoding canonical pathways for catabolism of catechol, benzoate, protocatechuate, phenylacetate, and methylated aromatic compound (Davis et al. 2012). *A. methanolica* 239 (previously known as *Streptomyces* sp. strain 239 or as *Nocardia* sp. strain 239) can grow in mineral medium broth containing methanol, ethanol, 1-propanol, 1-butanol, 2,3-butanediol, acetone, benzoic acid methylester, benzylamine, 3- and 4-hydroxybenzoates, 3,4-dihydroxybenzoate, phenylacetate, phenylacetaldehyde, phenyllactate, phenylpyruvate, 4-hydroxyphenylacetate, 4-hydroxyphenylpyruvate, D-phenylalanine, gentisate, and homogentisate as sole sources of carbon. Also, it contains

degradation pathways for benzoate, fluorobenzoate, toluene, xylene, styrene, naphthalene and other related polycyclic aromatic hydrocarbons (Wattam et al. 2014). *A. thermoflava* N1165 has been shown to degrade atrazine, naphthalene, anthracene, tetrachloroethene, 1- and 2-methylnaphthalene, 2,4-dichlorobenzoate, toluene, xylene, biphenyl, hexachlorocyclohexane, trinitrotoluene, ethylbenzene, and styrene (Chun et al. 1999).

Finally, cultivation and mass spectrometry evidences are provided that demonstrated that the catabolome array can aid in the understanding of degrading capacities without previous genome, and possibly metagenome, sequence knowledge.

Acknowledgement This work was supported by the CSIC (Programa CSIC de Cooperación Científica para el Desarrollo i-COOP; ref. COOPB20077), the European Community project KILLSPILL (FP7-KBBE-2012-312139), and the Spanish Ministry of Economy and Competitiveness (CTQ2014-55279-R). N.B. is a recipient of a fellowship from the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. We thank Drs. Dietmar H. Pieper and Ramiro Vilchez-Vargas for their support in providing resources for this study.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

Albarracín VH, Alonso-Vega P, Trujillo ME, Amoroso MJ, Abate CM (2010a) *Amycolatopsis tucumanensis* sp. nov., a copper-resistant actinobacterium isolated from polluted sediments. Int J Syst Evol Microbiol 60(2):397-401.

511 Albarracín VH, Amoroso MJ, Abate CM (2010b) Bioaugmentation of copper polluted soil
512 microcosms with *Amycolatopsis tucumanensis* to a copper for Zea mays plants.
513 Chemosphere 79(2):131-137.

514 Alcaide M, Tornés J, Stogios PJ, Xu X, Gertler C, Di Leo R, Bargiela R, Lafraya A, Guazzaroni
515 ME, López-Cortés N, Chernikova TN, Golyshina OV, Nechitaylo TY, Plumeier I, Pieper
516 DH, Yakimov MM, Savchenko A, Golyshin PN, Ferrer M (2013) Single residues dictate
517 the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family.
518 Biochem J 454(1):157-166.

519 Bailey MJ, Lilley AK, Thompson IP, Rainey PB, Ellis RJ (1995) Site directed chromosomal
520 marking of a fluorescent pseudomonad isolated from the phytosphere of sugar beet;
521 stability and potential for marker gene transfer. Mol Ecol 4(6):755-763.

522 Balachandran C, Duraipandiyan V, Balakrishna K, Ignacimuthu S (2012) Petroleum and
523 polycyclic aromatic hydrocarbons (PAHs) degradation and naphthalene metabolism in
524 *Streptomyces* sp. (ERI-CPDA-1) isolated from oil contaminated soil. Bioresour Technol
525 112:83-90.

526 Bargiela R, Gertler C, Magagnini M, Mapelli F, Chen J, Daffonchio D, Golyshin PN, Ferrer M
527 (2015) Degradation network reconstruction in uric acid and ammonium amendments in oil-
528 degrading marine microcosms guided by metagenomic data. Front Microbiol 6:1270.

529 Bargiela R, Mapelli F, Rojo D, Chouaia B, Tornés J, Borin S, Richter M, Del Pozo MV,
530 Cappello S, Gertler C, Genovese M, Denaro R, Martínez-Martínez M, Fodelianakis S,
531 Amer RA, Bigazzi D, Han X, Chen J, Chernikova TN, Golyshina OV, Mahjoubi M,
532 Jaouanil A, Benzha F, Magagnini M, Hussein E, Al-Horani F, Cherif A, Blaghen M,
533 Abdel-Fattah YR, Kalogerakis N, Barbas C, Malkawi HI, Golyshin PN, Yakimov MM,
534 Daffonchio D, Ferrer M (2015) Bacterial population and biodegradation potential in
535 chronically crude oil-contaminated marine sediments are strongly linked to temperature.
536 Sci Rep 5:11651.

537 Bastiaens L, Springael D, Wattiau P, Harms H, deWachter R, Verachtert H, Diels L (2000)
538 Isolation of adherent polycyclic aromatic hydrocarbon (PAH)-degrading bacteria using
539 PAH-sorbing carriers. Appl Environ Microbiol 66(5):1834-1843.

540 Beazley MJ, Martinez RJ, Rajan S, Powell J, Piceno YM, Tom LM, Andersen GL, Hazen TC,
541 Van Nostrand JD, Zhou J, Mortazavi B, Sobecky PA (2012) Microbial community analysis
542 of a coastal salt marsh affected by the Deepwater Horizon oil spill. PLoS One 7:e41305.

543 Bourguignon N, Fuentes MS, Benimeli CS, Cuozzo SA, Amoroso MJ (2014) Aerobic removal
544 of methoxychlor by a native *Streptomyces* strain: Identification of intermediate metabolites.
545 Int Biodeter Biodegr 96:80-86.

546 Bourguignon N, Isaac P, Alvarez H, Amoroso MJ, Ferrero MA (2014) Enhanced polyaromatic
547 hydrocarbon degradation by adapted cultures of actinomycete strains. J Basic Microbiol
548 54(12):1288-1294.

549 Brezna B, Khan AA, Cerniglia CE (2003) Molecular characterization of dioxygenases from
550 polycyclic aromatic hydrocarbon-degrading *Mycobacterium* spp. FEMS Microbiol Lett
551 223(2):177-183.

552 Chen BS, Otten LG1, Resch V1, Muyzer G2, Hanefeld U1 (2013) Draft genome sequence of
553 *Rhodococcus rhodochrous* strain ATCC 17895. Stand Genomic Sci 9(1):175-184.

554 Chun J, Kim SB, Oh YK, Seong CN,Lee DH, Bae KS, Lee KJ,Kang SO, Hah YC, Goodfellow
555 M (1999) *Amycolatopsis thermoflava* sp. nov., a novel soil actinomycete from Hainan
556 Island, China. Int J Syst Bacteriol 49:1369-1373.

557 Colin VL, Castro MF, Amoroso MJ, Villegas LB (2013) Production of bioemulsifiers by
558 *Amycolatopsis tucumanensis* DSM 45259 and their potential application in remediation
559 technologies for soils contaminated with hexavalent chromium. J Hazard Mater 261:577-
560 583.

561 Dávila Costa JS, Albarracín VH, Abate CM (2011a) Cupric reductase activity in copper-
562 resistant *Amycolatopsis tucumanensis*. Water Air Soil Pollut 216:527-535.

563 Dávila Costa JS, Albarracín VH, Abate CM (2011b) Responses of environmental *Amycolatopsis*
564 strains to copper stress. Ecotoxicol Environ Saf 74(7):2020-2028.

565 Dávila Costa JS, Kothe E, Abate CM, Amoroso MJ (2012) Unraveling the *Amycolatopsis*
566 *tucumanensis* copper-resistome. Biometals 25(5):905-917.

567 Davis JR, Goodwin LA, Woyke T, Teshima H, Bruce D, Detter C, Tapia R, Han S, Han J,
568 Pitluck S, Nolan M, Mikhailova N, Land ML, Sello JK (2012) Genome sequence of
569 *Amycolatopsis* sp. strain ATCC 39116, a plant biomass-degrading actinomycete. J
570 Bacteriol 194(9):2396-2397.

571 DeBruyn JM, Mead TJ, Sayler GS. Horizontal transfer of PAH catabolism genes in
572 Mycobacterium: evidence from comparative genomics and isolated pyrene-degrading
573 bacteria. Environ Sci Technol 46(1):99-106.

574 Duarte M, Jauregui R, Vilchez-Vargas R, Junca H, Pieper DH (2014) AromaDeg, a novel
575 database for phylogenomics of aerobic bacterial degradation of aromatics. Database
576 (Oxford) 2014:bau118.

577 Fathepure BZ (2014) Recent studies in microbial degradation of petroleum hydrocarbons in
578 hypersaline environments. Front Microbiol 5:173.

579 Guazzaroni M-E, Herbst FA, Lores I, Tamames J, Peláez AI, López-Cortés N, Alcaide M, Del
580 Pozo MV, Vieites JM, von Bergen M, Gallego JL, Bargiela R, López-López A, Pieper DH,
581 Rosselló-Móra R, Sánchez J, Seifert J, Ferrer M (2013) Metaproteogenomic insights
582 beyond bacterial response to naphthalene exposure and bio-stimulation. ISME J 7(1):122-
583 136.

584 Gutierrez T, Singleton DR, Berry D, Yang T, Aitken MD, Teske A (2013) Hydrocarbon-
585 degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation
586 and DNA-SIP. ISME J 7(11):2091-2104.

587 Hamzah RY, Al-Baharna BS (1994) Catechol ring-cleavage in *Pseudomonas cepacia*: the
588 simultaneous induction of ortho and meta pathways. Appl Microbiol Biotechnol 41(2):250-
589 256.

590 Hazen TC, Dubinsky EA, DeSantis TZ, Andersen GL, Piceno YM, Singh N, Jansson JK, Probst
591 A, Borglin SE, Fortney JL, Stringfellow WT, Bill M, Conrad ME, Tom LM, Chavarria KL,
592 Alusi TR, Lamendella R, Joyner DC, Spier C, Baelum J, Auer M, Zemla ML, Chakraborty
593 R, Sonnenthal EL, D'haeseleer P, Holman HY, Osman S, Lu Z, Van Nostrand JD, Deng Y,
594 Zhou J, Mason OU (2010) Deep-sea oil plume enriches indigenous oil-degrading bacteria.
595 Science 330(6001):204-208.

596 Iida T, Mukouzaka Y, Nakamura K, Kudo T (2002) Plasmid-borne genes code for an angular
597 dioxygenase involved in dibenzofuran degradation by *Terrabacter* sp. strain YK3. Appl
598 Environ Microbiol 68(8):3716-3723.

599 Kallimanis A, Karabika E, Mavromatis K, Lapidus A, Labutti KM, Liolios K, Ivanova N,
600 Goodwin L, Woyke T, Velentzas AD, Perisynakis A, Ouzounis CC, Kyrpides NC,
601 Koukkou AI, Drainas C (2011) Complete genome sequence of *Mycobacterium* sp. strain
602 (Spyr1) and reclassification to *Mycobacterium gilvum* Spyr1. Stand Genomic Sci 5(1):144-
603 153.

604 Kelley I, Cernigilia CE (1995) Degradation of a mixture of high-molecular-weight polycyclic
605 aromatic hydrocarbons by a *Mycobacterium* strain PYR-1. J Soil Contam 4(1):77-91.

606 Kim SJ, Kweon O, Freeman JP, Jones RC, Adjei MD, Jhoo JW, Edmondson RD, Cerniglia CE
607 (2006) Molecular cloning and expression of genes encoding a novel dioxygenase involved

608 in low- and high-molecular-weight polycyclic aromatic hydrocarbon degradation in
609 *Mycobacterium vanbaalenii* PYR-1. Appl Environ Microbiol 72(2):1045-1054.

610 Kim SJ, Kweon O, Jones RC, Edmondson RD, Cerniglia CE (2008) Genomic analysis of
611 polycyclic aromatic hydrocarbon degradation in *Mycobacterium vanbaalenii* PYR-1.
612 Biodegradation 19(6):859-881.

613 Kim SJ, Kweon O, Jones RC, Freeman JP, Edmondson RD, Cerniglia CE (2007) Complete and
614 integrated pyrene degradation pathway in *Mycobacterium vanbaalenii* PYR-1 based on
615 systems biology. J Bacteriol 189(2):464-472.

616 Kimes NE, Callaghan AV, Aktas DF, Smith WL, Sunner J, Golding B, Drozdowska M, Hazen
617 TC, Suflita JM, Morris PJ (2013) Metagenomic analysis and metabolite profiling of deep-
618 sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill. Front
619 Microbiol 4:50.

620 Kostka JE, Prakash O, Overholt WA, Green SJ, Freyer G, Canion A, Delgardio J, Norton N,
621 Hazen TC, Huettel M (2011) Hydrocarbon-degrading bacteria and the bacterial community
622 response in gulf of Mexico beach sands impacted by the deepwater horizon oil spill. Appl
623 Environ Microbiol 77(22):7962-7974.

624 Kubota M, Nodate M, Yasumoto-Hirose M, Uchiyama T, Kagami O, Shizuri Y, Misawa N
625 (2005) Isolation and functional analysis of cytochrome P450 CYP153A genes from various
626 environments. Biosci Biotechnol Biochem 69(12):2421-2430.

627 Kwak Y, Park GS1, Lee SE1, Li QX2, Shin JH3 (2014) Genome sequence of *Mycobacterium*
628 *aromaticivorans* JS19b1(T), a novel isolate from Hawaiian soil. J Biotechnol 186:137-138.

629 Kweon O, Kim SJ, Blom J, Kim SK, Kim BS, Baek DH, Park SI, Sutherland JB, Cerniglia CE
630 (2015) Comparative functional pan-genome analyses to build connections between
631 genomic dynamics and phenotypic evolution in polycyclic aromatic hydrocarbon
632 metabolism in the genus *Mycobacterium*. BMC Evol Biol 15:21.

633 Larkin MJ, Kulakov LA, Allen CC (2006) Biodegradation by members of the genus
634 *Rhodococcus*: biochemistry, physiology, and genetic adaptation. Adv Appl Microbiol 59:1-
635 29.

636 Liang Y, Van Nostrand JD, Deng Y, He Z, Wu L, Zhang X, Li G, Zhou J (2011) Functional
637 gene diversity of soil microbial communities from five oil-contaminated fields in China.
638 ISME J 5(3):403-413.

639 Lincoln SA, Hamilton TL, Valladares Juárez AG, Schedler M, Macalady JL, Müller R, Freeman
640 KH (2015) Draft genome sequence of the piezotolerant and crude oil-degrading bacterium
641 *Rhodococcus qingshengii* strain TUHH-12. Genome Announc 3(2) pii: e00268-15.

642 Lloyd-Jones G, Hunter DW (1997) Characterization of fluoranthene- and pyrene-degrading
643 *Mycobacterium*-like strains by RAPD and SSU sequencing. FEMS Microbiol Lett
644 153(1):51-56.

645 Lu Z, Deng Y, Van Nostrand JD, He Z, Voordeckers J, Zhou A, Lee YJ, Mason OU, Dubinsky
646 EA, Chavarria KL, Tom LM, Fortney JL, Lamendella R, Jansson JK, D'haeseleer P, Hazen
647 TC, Zhou J (2012) Microbial gene functions enriched in the Deepwater Horizon deep-sea
648 oil plume. ISME J 6(2):451-460.

649 Luo Q, Hiessl S, Steinbüchel A (2014) Functional diversity of *Nocardia* in metabolism. Environ
650 Microbiol 16(1):29-48.

651 Mason OU, Hazen TC, Borglin S, Chain PS, Dubinsky EA, Fortney JL, Han J, Holman HY,
652 Hultman J, Lamendella R, Mackelprang R, Malfatti S, Tom LM, Tringe SG, Woyke T,
653 Zhou J, Rubin EM, Jansson JK (2012) Metagenome, metatranscriptome and single-cell
654 sequencing reveal microbial response to Deepwater Horizon oil spill. ISME J 6(9):1715-
655 1727.

656 Mason OU, Scott NM, Gonzalez A, Robbins-Pianka A, Bælum J, Kimbrel J, Bouskill NJ,
657 Prestat E, Borglin S, Joyner DC, Fortney JL, Jurelevicius D, Stringfellow WT, Alvarez-

Cohen L, Hazen TC, Knight R, Gilbert JA, Jansson JK (2014) Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. ISME J 8(7):1464-75.

McLeod MP1, Warren RL, Hsiao WW, Araki N, Myhre M, Fernandes C, Miyazawa D, Wong W, Lillquist AL, Wang D, Dosanjh M, Hara H, Petrescu A, Morin RD, Yang G, Stott JM, Schein JE, Shin H, Smailus D, Siddiqui AS, Marra MA, Jones SJ, Holt R, Brinkman FS, Miyauchi K, Fukuda M, Davies JE, Mohn WW, Eltis LD (2006) The complete genome of *Rhodococcus* sp. RHA1 provides insights into a catabolic powerhouse. Proc Natl Acad Sci USA 103(42):15582-7.

Monticello DJ (2000) Biodesulfurization and the upgrading of petroleum distillates. Curr Opin Biotechnol 11(6):540-6.

Nozaki M, Kagamiyama H, Hayaishi O (1963) Metapyrocatechase. I. Purification, crystallization and some properties. Biochemische Z 338:582-590.

Pandey J, Chauhan A, Jain RK (2009) Integrative approaches for assessing the ecological sustainability of in situ bioremediation. FEMS Microbiol Rev 33(2):324-375.

Pathak A, Green SJ, Ogram A, Chauhan A (2013) Draft genome sequence of *Rhodococcus* opacus strain M213 shows a diverse catabolic potential. Genome Announc 1(1) pii: e00144-12.

Pérez-Pantoja D, De la Iglesia R, Pieper DH, González B (2008) Metabolic reconstruction of aromatic compounds degradation from the genome of the amazing pollutant-degrading bacterium *Cupriavidus necator* JMP134. FEMS Microbiol Rev 32(5):736-794.

Pérez-Pantoja D, Donoso R, Agulló L, Córdova M, Seeger M, Pieper DH, González B (2012) Genomic analysis of the potential for aromatic compounds biodegradation in Burkholderiales. Environ Microbiol 14(5):1091-1117.

682 Pizzul L, Castillo MP, Strenström J (2006) Characterization of selected actinomycetes
683 degrading polyaromatics hydrocarbons in liquid culture and spiked soil. World J Microbiol
684 Biotechnol 22(7):745-752.

685 Qu J, Miao LL, Liu Y, Liu ZP (2015) Complete genome sequence of *Rhodococcus* sp. Strain
686 IcdP1 shows diverse catabolic potential. Genome Announc 3(4) pii: e00711-15.

687 Sekine M., S. Tanikawa, S. Omata, M. Saito, T. Fujisawa, N. Tsukatani, T. Tajima, T.
688 Sekigawa, Kosugi H, Matsuo Y, Nishiko R, Imamura K, Ito M, Narita H, Tago S, Fujita N,
689 Harayama S (2006) Sequence analysis of three plasmids harboured in *Rhodococcus*
690 erythropolis strain PR4. Environ Microbiol 8(2):334–346.

691 Seo JS, Keum YS, Li QX (2009) Bacterial degradation of aromatic compounds. Int J Environ
692 Res Public Health 6(1):278-309.

693 Seo JS, Keum YS, Li QX (2012) *Mycobacterium aromativorans* JS19b1(T) degrades
694 phenanthrene through C-1,2, C-3,4 and C-9,10 dioxygenation pathways. Int Biodeterior
695 Biodegradation 70:96-103.

696 Shen FT, Lin JL, Huang CC, Ho YN, Arun AB, Young LS, Young CC. Shen FT, Lin JL, Huang
697 CC, Ho YN, Arun AB, Young LS, Young CC (2009) Molecular detection and phylogenetic
698 analysis of the catechol 1,2-dioxygenase gene from *Gordonia* spp. Syst Appl Microbiol
699 32(5):291-300.

700 Stingley RL, Brezna B, Khan AA, Cerniglia CE (2004) Novel organization of genes in a
701 phthalate degradation operon of *Mycobacterium vanbaalenii* PYR-1. Microbiology
702 150(11):3749-3761.

703 Sudhir KS, Hai G, Modi DR, Joyotsana KP (2014) Growth potential assessment of
704 actinomycetes isolated from petroleum contaminated soil. J Bioremed Biodeg 5:7

705 van Beilen J.B., R. Holtackers, D. Luscher, U. Bauer, B. Witholt, W.A. Duetz (2005)
706 Biocatalytic production of perillyl alcohol from limonene by using a novel *Mycobacterium*

sp. cytochrome P450 alkane hydroxylase expressed in *Pseudomonas putida*. Appl Environ Microbiol 71(4):1737–1744.

van Beilen JB, Funhoff EG, van Loon A, Just A, Kaysser L, Bouza M, Holtackers R, Röthlisberger M, Li Z, Witholt B (2006) Cytochrome P450 alkane hydroxylases of the CYP153 family are common in alkane-degrading eubacteria lacking integral membrane alkane hydroxylases. Appl Environ Microbiol 72(1):59-65.

Vila J, López Z, Sabaté J, Minguillón C, Solanas AM, Grifoll M (2001) Identification of a novel metabolite in the degradation of pyrene by *Mycobacterium* sp. strain AP1: actions of the isolate on two- and three-ring polycyclic aromatic hydrocarbons. Appl Environ Microbiol 67(12):5497-5505.

Vilchez-Vargas R, Geffers R, Suárez-Diez M, Conte I, Waliczek A, Kaser VS, Kralova M, Junca H, Pieper DH (2013) Analysis of the microbial gene landscape and transcriptome for aromatic pollutants and alkane degradation using a novel internally calibrated microarray system. Environ Microbiol 15(4):1016-1039.

Watanabe K, Hamamura N (2003) Molecular and physiological approaches to understanding the ecology of pollutant degradation. Curr Opin Biotechnol 14(3):289-295.

Wattam AR, Abraham D, Dalay O, Disz TL, Driscoll T, Gabbard JL, Gillespie JJ, Gough R, Hix D, Kenyon R, Machi D, Mao C, Nordberg EK, Olson R, Overbeek R, Pusch GD, Shukla M, Schulman J, Stevens RL, Sullivan DE, Vonstein V, Warren A, Will R, Wilson MJC, Seung Yoo H, Zhang C, Zhang Y, Sobral BW (2014). PATRIC, the bacterial bioinformatics database and analysis resource. Nucl Acids Res 42 (D1): D581-D59.

Whyte LG, Smits TH, Labbé D, Witholt B, Greer CW, van Beilen JB (2002) Gene cloning and characterization of multiple alkane hydroxylase systems in *Rhodococcus* strains Q15 and NRRL B-16531. Appl Environ Microbiol 68(12):5933-5942.

731 Xu L, Huang H, Wei W, Zhong Y, Tang B, Yuan H, Zhu L, Huang W, Ge M, Yang S, Zheng
732 H, Jiang W, Chen D, Zhao GP, Zhao W (2014) Complete genome sequence and
733 comparative genomic analyses of the vancomycin-producing *Amycolatopsis orientalis*.
734 BMC Genomics 15:363.

735 Yang X, Rui Xue, Chong Shen, Shuren Li, Chong Gao, Qi Wang, Xiaoxia Zhao (2011) Genome
736 sequence of *Rhodococcus* sp. strain R04, a polychlorinated-biphenyl biodegrader. J
737 Bacteriol 193(18):5032-5033.

738 Zeinali M, Vossoughi M, Ardestani SK (2007) Characterization of a moderate thermophilic
739 Nocardia species able to grow on polycyclic aromatic hydrocarbons. Lett Appl Microbiol
740 45(6):622-628.

741 Zhang C, Anderson AJ (2012) Polycyclic aromatic hydrocarbon degrading gene islands in five
742 pyrene-degrading *Mycobacterium* isolates from different geographic locations Can J
743 Microbiol 58(1):102-111.

744 Zhang Y, Qin F, Qiao J, Li G, Shen C, Huang T, Hu Z (2012) Draft genome sequence of
745 *Rhodococcus* sp. strain P14, a biodegrader of high-molecular-weight polycyclic aromatic
746 hydrocarbons. J Bacteriol 194(13):3546.

747 Zhao K, Guo X, Gong J (2013) A novel benzoate-degrading *Rhodococcus* strain contains three
748 *cat*A genes with one being transcriptionally active during the growth on benzoate. J
749 Environ Biol 34(2):401-407.

Table 1 List of probes associated to the 4 catabolic genes families that were found to be targeted by the microarray in *A. tucumanensis* DSM 45259.

Probe ID	Fluorescence (signal-background)[1]			Fold Change[1]	AromaDeg annotation[2]
ALKB_AJ833983	5360	6022	5757	43,634	Alkane hydroxylase (AlkB)
ALKB_AJ833926	2655	2781	2220	21,416	
RHDO_AAD17377	1968	3694	3733	20,655	Benzoate dioxygenase (Bzt)
RHDO_AAD03558	361	772	1433	6,302	Phthalate 4,5-dioxygenase (Pht)
RHMO_Z36909	6173	6054	5585	44,98	Phenol hydroxylase (PH)

[1]Signal corresponding to triplicates with standard deviation shown in Supplementary Table S1. Average (for triplicates) fold change of probe signal compared to background signal in the microarray is shown. [2]AromaDeg-based annotations obtained when each query sequence from probes targeting catabolic genes for which a positive signal was obtained in the microarray was submitted to web-based AromaDeg resource (Duarte et al. 2014).

Table 2 List of chemical signatures of key metabolites known to participate in the degradation of alkanes and aromatics in culture and control microcosms.

	Abundance (a.u.)[1,2]			
	0 h	24 h	48 h	72 h
n-Decanoic acid[3]	0	2071359	2736496	4518090
n-Tetradecanoic acid[4]	0	2458412	3202831	4382836
Phthalic acid[5]	79095538	64001448	52624286	39758652
Protocatechuic acid[5]	0	0	770651	1112138
3-Oxoadipic acid[5]	4110	26955989	40657623	25657381
Phenol[6]	862601	498098	196556	93768
Catechol[6]	165585	60114743	71964381	98848577
Benzoate[7]	1874035	1376927	1106361	794551
Catechol[7]	532343	34588533	49513068	58615247

[1]Abundance (in arbitrary units) was calculated (in triplicates) as the area of the peak (calculated on the basis of *m*/*z* and/or standards) of chemicals determined by LC-MS (positive [+] and negative [-] polarities) in cultures containing the selected pollutants. Strain DSM 45259 was cultivated on minimal mineral medium with *n*-decane, phenol, benzoate and phthalate (500 mg L^{-1}) as the sole source of carbon and energy, at 30 ºC and 180 rpm during 72 h. Quantification of the biodegradation was further performed by extraction and target analysis of the substrates by LC-MS. Detailed conditions for cultivation and analytics are given in Materials and Methods section. [2]Abundance levels for biotic and abiotic controls were considered for background corrections. [3]Abundance levels of the initial pollutant and degradation intermediates in cultures with *n*-decane. [4]Abundance levels of the initial pollutant and degradation intermediates in cultures with *n*-tetradecane. [5]Abundance levels of the initial pollutant and degradation intermediates in cultures with phthalic acid; the presence of small amount of oxoadipic acid may be due to the presence of small amount of cells added at the beginning of the assay (see Materials and Methods section for details). [6]Abundance levels of the initial pollutant and degradation intermediates in cultures with phenol. [7]Abundance levels of the initial pollutant and degradation intermediates in cultures with benzoate.

Figure legends

Figure 1 Potential alkane and aromatic catabolic abilities of *A. tucumanensis* DSM 45259 guided by microarray data. Solid lines represent single step reactions while dotted lines represent degradation steps where multiple reactions are involved. Enzyme codes as follows: Alkane hydroxylase (AlkB); Benzoate dioxygenase (Bzt); Phthalate 4,5-dioxygenase (Pht); Phenol hydroxylase (PH). Red color names indicated enzymes encoded by genes targeting probes in the catabolome microarray, whereas those with blue color indicates those whose presence was unambiguously demonstrated by the presence of degradation intermediates (see also molecules in blue color) formed by the action of such enzymes. As shown in Table 2, the presence of *n*-decanoic acid, catechol, protocatechuate and 3-oxoadipate was confirmed by target LC-MS analysis. The degradation of catechol was confirmed by measurement of catechol-2,3-dioxygenase activity. Note: *n*-decane has been used as example of the ability of strain DSM 45259 to degrade alkanes, although cultivation tests and target metabolomics analysis also demonstrated its ability to degrade *n*-tetradecane.